

DOCUMENT RESUME

ED 268 146

TM 860 072

AUTHOR Steele, Joe M.
TITLE Trends and Patterns in Writing Assessment.
PUB DATE 8 Mar 85
NOTE 8p.; Paper presented at the Annual Conference on the Assessment of Writing (3rd, San Francisco, CA, March 8, 1985).
PUB TYPE Speeches/Conference Papers (150) -- Reports - General (140)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cognitive Processes; *Educational Trends; *Essay Tests; Evaluation Criteria; *Evaluation Methods; *Evaluation Problems; Higher Education; *Persuasive Discourse; Student Evaluation; Verbal Communication; *Writing Evaluation; Writing Instruction; Writing Skills
IDENTIFIERS *College Outcome Measures Project

ABSTRACT

Various approaches have traditionally been used to evaluate college students' writing performance. The College Outcome Measures Program (COMP), designed to assess and improve college programs, offers new ways to collect writing samples that include responses to reasoning and problem solving tasks. Thus, the effectiveness of written communication can be assessed, as well as the mechanics. Many writing samples have involved abstract essays addressed to an unspecified audience, rather than tasks which represent the kind of writing commonly done outside the classroom--such as letters which must make contact with the reader in a convincing way. Traditionally, students' writing has been evaluated to judge technical accuracy and mechanics, and to assign a grade, rather than to provide a reaction to students' ideas and to help in clarifying their thinking. Ironically, although evaluation threatens students, it is necessary to foster excellence. In COMP writing exercises, students respond to situational letter-writing tasks, and are provided with feedback on the plausibility and appropriateness of their ideas. Rather than one holistic rating, COMP defines proficiency levels on: (1) making contact with the relevant audience; (2) organizing ideas; and (3) using language in a clear and interesting way. At various points of instruction, COMP obtains writing samples in three content areas in order to achieve a reliable and valid measure of a student's level of proficiency. (GDC)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED268146

TRENDS AND PATTERNS IN WRITING ASSESSMENT

Joe M. Steele

Presented at the Third Annual
Conference on the Assessment of Writing
San Francisco, California
March 8, 1985

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

J. M. Steele

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

COLLEGE OUTCOME MEASURES PROGRAM

AMERICAN COLLEGE TESTING PROGRAM

IOWA CITY, IOWA 52243

TM 860072

TRENDS AND PATTERNS IN WRITING ASSESSMENT

Joe M. Steele
American College Testing Program

The trends in writing assessment give rise to certain patterns and concerns that are the focus of this presentation. We are the assessed society. As Arthur Appleby points out, much or most of what students write will be assessed in some fashion. What is this written material that we are judging? How do we as faculty and administrators judge it? How do we use it? What implications do the answers to these questions have for the quality of instruction and performance in writing? My comments in answer to these questions arise out of observations in working with hundreds of colleges and universities that are participating in the College Outcome Measures Program (COMP). COMP, as an assessment of general education outcomes of college, is primarily a range of services and assessments to help institutions assess and improve the effectiveness of their programs. In this context, COMP offers ways to collect writing and speaking samples that include responses to reasoning/problem solving tasks. Since COMP assessments offer a contrast to current trends in writing assessment, I will return to our approach later.

Student Responses: What Are We Judging?

The first, and perhaps, key feature of writing samples is their meaningfulness as the raw material for making judgments of skill levels. Many topics elicit formal or abstract essays addressed to a general audience (or no one in particular) rather than tasks that represent the kind of writing commonly done outside the classroom, such as letters to friends, representatives, or community leaders on topics that represent public/social issues. Often the only criteria implied for judging the writing are editorial/technical accuracy and mechanics rather than a concern about making contact with the reader and communicating ideas in a convincing way. Students as writers have learned that they are

asked to write so that teachers can sort them into groups or assign a grade rather than receive a reaction to their ideas and help in clarifying their thinking.

The way students have been involved as participants will certainly affect their attitude and response to writing tasks. But even if good natured, the response of some (or many) students may be half-hearted. They may, after years of being pinched and measured, not take the assessment seriously or may lack the willingness, commitment and energy to do their best work. In the context of a required writing class, many may not be motivated to try to improve their writing skills; they may not see the relevance of the class or the writing assessment. Some will simply be playing the student "game" of trying to second guess (or outwit) the teacher.

It appears to be the case that writing assessments are frequently used without explaining the criteria for judging and with no feedback and explanation of results. For the COMP assessments of writing, we feel students should at a minimum be told the following:

"The writing assessment asks you to role-play three situations in which you write a letter. You should try to put yourself in each situation and respond to the best of your ability as if you were actually in that adult role. We will rate your ability to write a persuasive message to the specified audience, developing as many ideas as you can. We will rate the way you organize your points and your use of correct and lively language to support your argument."

In addition, students should be offered the opportunity for feedback of results and some interpretation of their performance. Discussion of the plausibility and appropriateness of ideas generated could assist students in reasoning skills. In the context of a writing course, some follow-up activities related to audience contact, organization of ideas, and use of language might be appropriate.

Parallels in Faculty Responses

Faculty involvement sometimes parallels student involvement in the assessment of writing. Some faculty do not value the measures used and what they assess. This attitude is communicated to students in many ways and undermines the effort to obtain the best efforts of students. Thus, the way both students and faculty approach the use of writing assessments could undermine their validity.

Some faculty are interested only in the sorting function of the writing assessment. Results are neither fed back to students nor interpreted and used in instruction. The reasons for such obscurity may lie in concerns about either test security, teaching to the test, or sheer thoughtlessness. Moreover, if students may not care what they write and do not want to talk about it, faculty may not care what students have written and only wish to focus on the way they wrote it. This brings us to a basic paradox of evaluation. Gene Glass phrased the paradox this way: Excellence cannot prosper if evaluated; excellence cannot prosper unless evaluated! The paradox revolves around evaluation as an external threat and judgment versus evaluation as part of a shared judgment in a counseling or mentoring role where it is used to provide clarification and identify an understanding of relevant criteria. Rex Brown, in the opening address for this conference referred to this paradox in terms of engaged versus disengaged evaluation. Too often "evaluation" is merely equated with higher test scores.

The focus of much assessment of writing thus appears to be an external judging of the student—who often even has to pay the cost. This labelling of student performance often proceeds before the quality and effectiveness of the writing program has been clearly established. Instead of judging the student, data obtained from assessments of writing should be used to examine some of the following questions:

Does the program produce substantial growth in writing proficiency?

Does the program bring all students up to a specified level of proficiency?

Are sections of freshmen courses of even quality and emphasis?

Do the series of courses in a program evidence a well-defined sequence and integration?

Do the needs of weaker students eclipse efforts to improve skills of stronger students?

To support examination of such questions, COMP provides reference group norms in writing for freshmen and seniors at a number of institutions. Alternative, equated forms of the assessment are available. Currently, COMP can provide distributions of mean gains in overall growth in general education skills and knowledge for samples of seniors at many institutions. As more institutions utilize the COMP writing assessments, similar estimates of average gains in writing may be developed.

The Content of Assessment

What is the content of assessment in writing? First of all, content can refer to the nature of the task itself. Is the topic engaging and challenging? Does it call for a meaningful, purposeful kind of writing (rather than being merely a writing "exercise")? There is much debate about whether tasks should demand description, persuasion, or some other type of writing. COMP has taken the position that tasks should reflect the kind of writing demands that adults out of school encounter from day to day. Typical writing tasks often involve a letter in which one takes a position and attempts to persuade the recipient to support this position.

In designing the COMP writing assessments as part of a larger assessment of general education outcomes of college, a number of other elements are embedded in them. Each task calls for skills in problem solving and clarifying values. Thus, each task can be assessed for reasoning skills as well as proficiency in communicating in writing.

The content of assessment includes the criteria and approach used for judging the product as well as the topic and context that elicits the product. Holistic ratings that are commonly used for rating writing samples limit the way such judgments can be

used. First, they are not diagnostic. They do not identify strengths and weaknesses in the student's writing. Second, they often do not define criteria that are used as a basis for assigning a rating. Instead they tend to be norm-referenced, comparing a product with other papers in a particular set that are viewed as better or weaker. Multiple criteria, clearly defined, offer more help to both students and teachers. COMP has defined five levels of proficiency on three broad criteria: 1) Making and sustaining contact with the relevant audience, 2) Organizing ideas to address and develop the points called for, and 3) utilizing language and sentence structuring techniques to present the ideas in a clear and interesting way. Separate scales assess the ability to identify reasonable and appropriate benefits, problems and solutions related to the point of view called for in the role-playing task. Currently, a study is underway at ACT to examine the relationship between these writing and reasoning skills. A preliminary study indicates a Pearson product-moment correlation of about .50! This suggests only a modest relationship between the ability to generate a number of relevant ideas and the ability to structure an effective communication. The study will examine entering and exiting levels of proficiency in reasoning as well as writing. The study will also seek to identify programs that may be more effective in developing critical thinking skills.

In addition to the value and utility of multiple criteria for assessing writing, there is also a need for multiple writing samples to adequately judge students' current levels as well as growth in writing skills. There appears to be a tendency to obtain only one sample of writing from a student. Studies at both ACT and CSE have suggested that two or three writing samples are needed to obtain a sufficiently reliable and valid measure of a student's level of proficiency. The COMP writing assessment obtains three writing samples, structured to be content fair by including a topic in the social sciences, natural/biological sciences, and arts/humanities areas. By measuring entering levels of proficiency and sampling writing skills at various points of instruction, a portfolio of writing samples could be developed that would be extremely useful for both instructional

EAGLE-A

and program evaluation purposes.

Thus, trends in assessment of writing have caused many people to focus on the assessment process rather than on the student and the act of thinking. Concerns for efficiency and reliability seem to have caused many to choose tasks and scales that require the least time and training rather than yield the richest and most accurate information. Concerns for processing and sorting students sometimes appear to have overridden concerns for utility and effectiveness in instruction. Finally, concerns about writing skills have shifted attention (and perhaps efforts) away from the development of reasoning skills that ultimately make the writing worthwhile. This does not mean that we should abandon our efforts to better assess writing. It does suggest that we should select and perfect our practices as well as our assessment instruments to correct these imbalances.